WDQI Part II: Pennsylvania's *Data for Education and Workforce (PA-DEW)*

Introduction

The Pennsylvania Department of Labor & Industry (L&I) is pleased to submit its Workforce Data Quality Initiative proposal to the U.S. Department of Labor for funding to develop a longitudinal data system in close collaboration with the Pennsylvania Department of Education (PDE) that will yield critical information and analysis pertaining to workforce and education data. The Commonwealth of Pennsylvania is submitting this application as a single state application. In order to maximize this federal investment, the commonwealth will leverage several multi-state data sharing agreements.

To increase employability of Pennsylvanians, reduce unemployment, and generate sustainable economic growth, workforce developers and educators must be fully aware of the jobs and skills that are in high demand in the state's economy. Additionally, a clear understanding of the return on educational investments is essential to ensure that state and federal resources are aligned with existing and emerging industry needs. Educators and workforce development stakeholders benefit from being empowered to compare workforce performance data with educational programs and training delivered to Pennsylvanians who enter the workforce and/or pursue career-advancement. With a labor force in excess of 6 million, an estimated 834,000 student-pipeline (including traditional and non-traditional students), 270,000 businesses, 267 Institutions of Higher Education (IHE) and over 800 Local Education Agencies (LEAs), Pennsylvania is in critical need of a mechanism to generate data-intelligence that can be utilized to bridge existing skills gaps and continually prepare labor force participants to meet the needs of employers. Analysis and information generated from a longitudinal data system of workforce and education data will be the foundation that will enable the commonwealth to achieve this milestone. The

longitudinal data system will also be critical in terms of developing a continuous culture of partnership between educators, educational institutions, workforce and economic development stakeholders, and business leaders to jointly evaluate program effectiveness. Such a culture will ultimately lead to continuous program improvements, and when appropriate, new program development where gaps are identified. With sophisticated output and analysis possible from such a system, policy makers, educators, businesses, workforce and economic developers will be in a position to create new strategies that empower students, job seekers, and employers to make informed career and business decisions. The ultimate goal for Pennsylvania in creating this longitudinal data structure is to have a system that helps create and retain a skilled workforce that is responsive to industry needs and capable of driving economic growth.

Plan Overview: Current Capacity

Pennsylvania does not have an established education to workforce longitudinal data system. L&I workforce data is housed in three main systems: the Unemployment Compensation Modernization System (UCMS), the Commonwealth Workforce Development System (CWDS) and the Data Mart (DM). UCMS contains unemployment compensation (UC) claims and payments data. CWDS contains information on unemployment compensation claimants and other registrants, including their job histories, services and training received under Workforce Investment Act (WIA), Wagner-Peyser, Veterans, National Emergency Grants, Trade, Office of Vocational Rehabilitation (OVR) and Department of Public Welfare (DPW) programs. The DM contains ten years of new hires information, unemployment compensation covered employee wage records, employment and wage data for employers and industry classification for employers. Data exchanges between these three systems are conducted both routinely for established reports and measures, as well as on an ad-hoc basis for special reports and studies.

Education data is housed in PDE's statewide longitudinal data system (SLDS), called the Pennsylvania Information Management System (PIMS). PIMS is a secure data collection and reporting system that allows educators to evaluate the educational progress of students from early childhood through higher education. Data are collected from a variety of Pennsylvania educational institutions. At the K-12 level this includes public school districts, Intermediate Units, charter schools, career and technical education centers (CTCs), the State Juvenile Corrections Institutions, and a small subset of private schools. At the postsecondary level, this includes the commonwealth's two-year community colleges and other Perkins-funded postsecondary institutions. Data elements collected on individuals are required by state and federal mandates and include a wide variety of variables, such as demographics, school attended, Pennsylvania System of School Assessment scores (PSSAs), full-time or part-time enrollment status, and type of student (elementary, high school, college, Perkins, adult learner, CTC, etc.). For postsecondary institutions, additional data (also required to be collected pursuant to state and federal mandates) includes degree awarded, CIP code (Classification of Instructional Program), and course credits earned. PIMS is currently used for established reports and measures and on an ad-hoc basis for special reports and studies.

To bridge the disconnect between workforce and education data systems, Pennsylvania seeks to build a federated data system, the Pennsylvania *Data for Education and Workforce* (PA-DEW) system, which will allow for analysis that ties education data to workforce outcomes in a combination of aptitude and programs. The analysis of program effectiveness will lead to a determination of whether a program needs to be modified to meet the demands of industry and the learners in that particular field of study. Through PA-DEW, more informed decisions will be made on local, regional and statewide levels that will ultimately increase the employability of

students exiting educational institutions and entering the workforce. Improved data directing job seekers to sustainable career paths will lead to more productive citizens who contribute to the state's economy and tax base. In the absence of PA-DEW, educational institutions cannot meaningfully or successfully collaborate with workforce and economic development practitioners and business leaders to crosswalk program offerings with industry demand.

1.1 Security

L&I's Office of Information Technology (OIT) maintains rigorous security measures for existing workforce datasets. For decades, OIT has been successful in maintaining secure wage records and unemployment insurance claims and payments data systems. These systems span several data platforms and technologies. Internal and external auditing ensures that access is granted on a least-privilege basis, and that proper maintenance and change controls are applied. L&I's OIT is experienced with data integration and how best to design and utilize the infrastructure and software for the creation of a federated data system that will enable the secure sharing of confidential data among state agencies, specifically between L&I and PDE. A federated data model has already proven successful in Pennsylvania through implementation of linkages between K-12 student-level data in PDE's PIMS with early learning data stored in the DPW Data Warehouse. This experience and success will ensure the integrity of the state's development of the PA-DEW.

The federated data system contains sophisticated programs that will combine the agencies confidential data through specified linkages. Using a de-identified model, various data elements will be combined together with safeguards to prevent the disclosure of personally identifiable information or identification of individuals. This protects privacy interests and complies with state and federal confidentiality requirements. The federated system approach allows partner

agencies to maintain the security of their own separate databases thereby limiting the possibility of users gaining unauthorized access.

Pennsylvania requires commonwealth-wide compliance governed by its Office of Administration with a mature state policy that maximizes security. In addition, L&I and PDE have strict oversight from agency legal counsel. A *Commonwealth Employee User Agreement* is signed by each commonwealth staff member before access is granted at his or her workstation. A similar document exists for non-commonwealth employees. Policy for *Data in Transit and Data at Rest* also requires each agency to evaluate data, determine its security requirement, and apply security where appropriate. Finally, individual program *Non-Disclosure Agreements* are signed by all staff in order to maintain regulatory compliance.
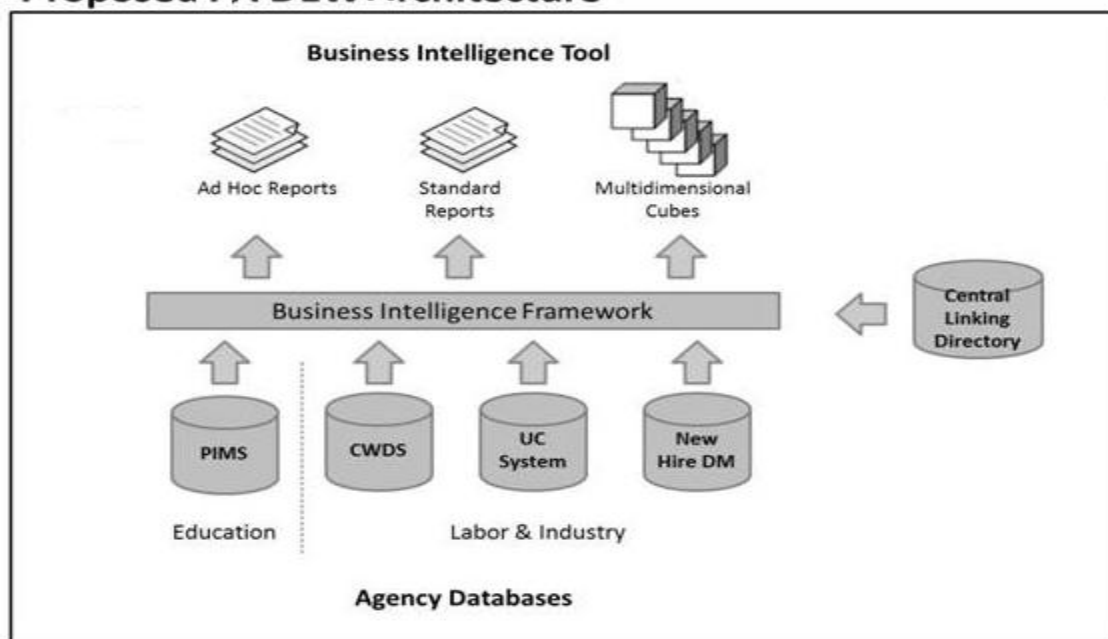
1.2 Education Partnerships

Through this grant, the commonwealth will link the education data housed in PDE's PIMS SLDS with the workforce data housed in L&I's UCMS, CWDS, and DM systems. Establishing this new federated data system fulfills Pennsylvania's long-term goal of conducting performance reporting and research on the education-workforce dataset assimilation.

1.3 Data Linkages

Through a federated data system, L&I's data will be linked to PDE's education data, which is housed in the PIMS SLDS. PDE has previously applied for and received funding through the NCES (National Center for Education Statistics) IES (Institute of Education Sciences) SLDS grant program to explore the linkage between education and workforce data. PDE and L&I are currently working together and are in the process of conducting the work outlined in the SLDS grant to determine the best approach to link the data. The findings of this work will be used to leverage the work in this grant to implement the linkage and develop the proposed architecture as

outlined below:

Proposed PA-DEW Architecture

1.4 Analytical Results from Partnerships

PA-DEW will allow for analysis of a number of aspects of the education and workforce development systems including:

For Students: Current methods of evaluating student outcomes after graduating from a postsecondary institution rely primarily on post-graduate surveys with inconsistent response rates. Leveraging linkages provided through this project, the commonwealth will be able to more effectively determine an individual's employment status following graduation. The ability to link industries of employment to programs of study will also enable Pennsylvania to better align education funding with industry needs. Specific measures available via PA-DEW include:

- *Number of students working in their field of study following graduation* to determine if programs are preparing students for work in their chosen fields (industry).

- *Comparison of wages for adults pre-college and post-graduation* to provide data for determining the impact of completion of a postsecondary credential on wage progression.

- *Comparison of wages earned by individuals with a postsecondary credential and those with only a high school diploma* to provide data illustrating the benefit that can be attributed to a postsecondary credential.

- *Type and location of industries that employ students with postsecondary credentials to allow* for the evaluation of relocation patterns within Pennsylvania by industry. Identifying industries that are drawing postsecondary graduates outside of their region can also be used to target areas of the commonwealth with workforce/postsecondary training needs and to identify postsecondary institutions to meet those needs. This analysis, developed by L&I in anticipation of PA-DEW, will take into account individuals finding employment outside Pennsylvania; a practice that can be replicated by other states.

- *Length of time between graduation and employment in a student's field of study* to provide information on market saturation for education fields.

For working students: Data linkages created through PA-DEW will allow for the identification of the number of students working while attending postsecondary institutions of higher education. Studies indicate that educational success is impacted by the number of hours worked while enrolled in postsecondary education. The ability to compare the success of students who worked either full or part time while enrolled in postsecondary education to the success of students who did not work while enrolled in postsecondary education will help institutions better design programs to serve working students.

For partial postsecondary completers: Since enrollment in postsecondary education is counter-cyclical to the state of the economy, postsecondary enrollment typically increases during times of high unemployment. Many of these students then return to the workforce when the

economy improves without completion of credential. According to the Council for Adult and Experiential Learning, almost 20 percent of Pennsylvania's workforce between ages 18 and 64 have some college education but no degree. Through PA-DEW, the commonwealth will be able to study the trends of these atypical student cohorts that leave school to re-join the workforce. Specific measures available include:

- *Number/percent of students who leave postsecondary education to accept employment and the types of industries in which they are employed;*

- *Employment rate for non-graduates who leave postsecondary education before earning a credential and the industries in which they find employment; and*

- *Type of postsecondary education sought by students who enrolled in a program leading to a credential in the industry where they were previously or are currently employed.*

For Adult Basic Education Students: The Workforce Investment Act (WIA) of 1998 requires Pennsylvania to measure the placement and retention rates of adult literacy students in unsubsidized employment, as well as educational gains, transition to postsecondary education, and obtainment of a secondary school credential. Since 2001, PDE and L&I have conducted a data match in Unemployment Insurance (UI) wage records for adult literacy students who either were employed upon enrollment in an adult literacy program or who set the goal of obtaining employment. The current process is cumbersome; but efficiency will be accomplished with the creation of PA-DEW. Other measures that will help PDE, L&I, and the state and Local Workforce Investment Boards (LWIBs) determine the return on investment of state and federal funds in adult literacy programs include: w*age increases as a result of program participation* and industries of employment for program completers, e*mployment retention rates* for adult

literacy participants, and *number of participants in adult literacy programs that go on to further training* through workforce programs such as WIA Title I and OVR.

For Career and Technical Education (CTE): Federal Perkins IV legislation requires Pennsylvania to identify employment status during the second quarter after a student exits CTE. Currently, employment data requirements are met through the use of an online follow-up survey application. Through PA-DEW, the commonwealth will be able to efficiently measure *employment status and wages earned* of students after exiting CTE, *retention rates and wage data* over time (not just the second quarter after exiting CTE), and *identification of high-demand, high-skill and high-wage jobs in relation to setting goals for CTE programs* to meet future labor market demands.

For Workforce Program Participants: In addition to evaluating the relationship of education to the workforce, an analysis of those enrolled in workforce programs compared to their education will be conducted. Information will be generated to show if there are specific educational programs that may not be preparing individuals for the workforce and what workforce training programs are needed to supplement the education received in those programs. This analysis will examine individuals who enter the workforce system and then move on to formal education at a postsecondary school. Employment outcome data will be tabulated under both scenarios to determine the effectiveness of the workforce system in providing appropriate services/training.

For Worker Supply: The supply of workers being prepared for an occupation will be analyzed by PA-DEW to provide a more accurate analysis of the educational attainment levels of participants in the workforce system (analysis to be utilized in regional supply/demand analysis described in section 3.2). One example of analytical enhancement includes analysis to determine

if employment outcomes are being negatively impacted by over-supply. Additionally, educational attainment level is currently self-reported. Levels reported through this initiative will aid in better understanding and interpreting participant outcomes.

1.5 Partnerships with Neighboring States

L&I is a signatory of the Wage Record Interchange System 2 (WRIS2) agreement which will enable the commonwealth to capture successful employment placements outside of Pennsylvania. In addition, L&I is in a unique position of crafting the *Interstate Data Sharing Agreement* (IDSA) with the Mid-Atlantic Region (Maryland, Washington D.C., Delaware, West Virginia, and Virginia) and federal agencies, including the U.S Departments of Education (USDE) and Labor (USDOL). The IDSA will complement Pennsylvania's WRIS2 agreement to strengthen partnerships with neighboring states to accurately monitor workforce outcomes in the full labor market. Although these data sets will not be integrated into PA-DEW in its initial stage, this future enhancement is considered critical by the commonwealth.

Section 2: Plan Outline

L&I and PDE will collaborate to develop PA-DEW to serve as the state's virtual workforce longitudinal database. The PA-DEW will link UCMS, CWDS, DM, and PIMS, all of which were described previously.

Section 2.1 Objectives for Data System Improvement

The commonwealth has been interested in linking education and workforce data and has already undertaken efforts to explore how best to create these linkages. In preparation for this WDQI submission, focused meetings between the two agencies (PDE and L&I) along with the Governor's Policy Office have resulted in clear definitions of challenges and identification of how to convert those challenges into opportunities. This collaboration between PDE and L&I,

guided by the Policy Office, has raised the spirit of cooperation and enhanced communication between the two departments.

PA-DEW will fill the void of connectivity between education and workforce information by integrating data from the respective systems and creating a virtual workforce longitudinal database within a federated system. The creation of PA-DEW is the first step in the commonwealth's long term vision of releasing the power of data integration. Currently, Pennsylvania's data resides in various warehouses, only some of which can be linked. Therefore, the first goal is to improve the quality of Pennsylvania's data systems by providing greater connectivity between workforce and education data. A secondary goal is to enhance connectivity within the three workforce data systems housed within L&I. The final and most crucial goal of this proposal is the analysis that will be possible as a result of the linkage between the education and workforce systems. This analysis will inform decision makers about how well our workforce and education programs are performing and will guide decisions, strategies and resource allocation moving forward.

2.2 Status of Longitudinal Education Data System

PIMS was launched in 2007 and is the PDE PreK-16 unit-level statewide longitudinal database. PIMS is a secure data collection and reporting system that is designed to meet state and federal reporting requirements and provide information to improve student achievement through more efficient and effective use of data. PIMS adheres to all Family Educational Rights and Privacy Act (FERPA) requirements and uses unique, anonymous student identifiers (PAsecureID) in order to protect a student's PII. In addition, PDE has implemented a *Data Access and Use* policy. This FERPA compliant policy guides staff who have access to student level data and prescribes how data in PDE student unit record systems will be collected,

maintained and disseminated in compliance with applicable federal and state laws. This policy

applies to all organizational units and staff within PDE, to LEAs and staff, to institutions of

postsecondary education (IPEs) and staff, to authorized agents of PDE, LEAs or IPEs and any

contractors, subcontractors and their agents, and to other individuals, agencies, and organizations

seeking to access or to utilize education records stored in PDE student unit record systems.

Additional information regarding the policy can be found on the PDE website at:

http://www.portal.state.pa.us/portal/server.pt/document/1224411/pims_data_access_policy_pdf_

%282%29?qid=50141450&rank=2.

PIMS collects detailed information from the educational entities. Data elements collected on

individuals are those that are required by state and federal mandates. At the K-12 level data is

collected from: public school districts, Intermediate Units, charter schools, career and technical

education centers (CTCs), the State Juvenile Corrections Institutions, and a small subset of

private schools. At the postsecondary level, data is collected from the commonwealth's two-year

community colleges and other Perkins-funded postsecondary institutions. Specific program data

collected or calculated in PIMS include: Student Enrollment, Special Education, Highly

Qualified Teacher, Professional Personnel, Subsidy Payments, Child Accounting, Safe Schools,

Cohort Graduation Rate, Career and Tech Education, Limited English Proficiency, Teacher Loan

Forgiveness, and Low Income.

In 2007, the USDE awarded PDE a $4 million grant to implement PIMS. By October 2007,

PIMS was collecting required data from over 800 K-12 educational entities throughout the

commonwealth. A second grant for $6.1 million from the USDE in 2009 allowed PDE to include

postsecondary data into PIMS. In 2009, DPW's Office of Child Development and Early

Learning also completed the process of linking data on students enrolled in publicly funded early

childhood education programs with PIMS data by assigning student identifiers to all students enrolled in these programs. Finally, in 2010, PDE was awarded a third USDE grant for $14.3 million to enhance PIMS. This third grant includes funding that has allowed PDE and L&I to begin exploring the task of linking the data so that the goals of this grant application can be accomplished. This third grant includes $182,000 in funding that has allowed PDE and L&I to begin exploring the task of linking the data and creating the infrastructure so that the goals of this grant application can be accomplished. These successes will be leveraged by PA-DEW.

2.3 Sustaining the Effort

The cost of the federated data system proposed by Pennsylvania is mostly incurred in its initial development. Ongoing costs after the grant period are minimal. In anticipation of applying for this grant, L&I convened meetings with past recipients of the WDQI to confirm this finding. States that Pennsylvania has conducted such conversations with include Virginia, Ohio, and Louisiana. After the grant has ended, ongoing cost of maintaining and operating the PA-DEW will be shared by various bureaus within L&I. Individual workforce and education agencies will continue to use existing resources to maintain and support their individual databases and to supply routine clean data extracts for PA-DEW as a result of the development work to be carried out during the life of the grant. L&I and PDE anticipate some labor savings from data standardization and from ad-hoc self-service reports. Currently, several PDE reports depend on costly surveys that are sent to program exiters. The PA-DEW will eventually eliminate the need for such laborious data-gathering mechanisms.

The governance structure proposed in this application will continue after the end of the grant to monitor data quality, address any ongoing issues regarding appropriate uses of the data system, and advocate for growth of the system as needed and as appropriate, including future

enhancements suggested in this proposal. However, at the most fundamental level, what will sustain this effort is that both L&I and PDE have worked together for many years to build a solid partnership and the commitment of the Secretary of Labor & Industry, Secretary of Education, and Governor's Policy Office to the project.

Section 3: Partnership Strategies

The WDQI will strengthen existing connections both internally in L&I and externally with PDE. The governance structure will include workforce agency partners, education partners, employers and economic development partners. The linkage to PDE's PIMS SLDS will allow analysis of the impact of education on labor force outcomes and on the use of public assistance, and how the role of educational attainment may break the cycle of unemployment and poverty.

3. 1 Partnerships with State Workforce System

Pennsylvania's state workforce system, led by the Secretary of L&I and the Governor's Workforce Investment Board (WIB), includes the 22 local WIBs, educators, policy makers, job seekers, employers, and numerous commonwealth agencies including the Departments of Education, Corrections, Community and Economic Development, Public Welfare, Aging, Health, Transportation, and Military and Veterans Affairs. L&I has strong ongoing relationships with all of these entities and will leverage their participation through the WIB. However, the multiple programs and bureaus within L&I whose functions are integrated to meet program mandates including UC, WIA, Trade, Veterans Programs, Vocational Rehabilitation, and Labor Market Information will be prioritized. All of these programs individually collect and maintain data on the customers they serve and the services they provide. With the WDQI grant opportunity, these operations will be integrated to enhance Pennsylvania's workforce system from a technological and operational perspective. Joint collaboration for the current development

of the UCMS, development and enhancements to the CWDS (one-stop), and the development

and maintenance of CWIA's DM, will be formalized as a first step in PA-DEW. Targeted focus

on connecting these systems will also expedite Pennsylvania's ability to incorporate educational

data into PA-DEW. PDE and L&I's experience in developing systems to account for reporting

requirements, as well as transactional activity, will serve well in developing PA-DEW. L&I's

expertise in producing workforce outcomes from administrative data, preparing supply/demand

analyses, and linking educational programs to occupations will enable Pennsylvania to hit the

ground running in conceptualizing and producing comprehensive reports assessing the value of

education, evaluating successes of the workforce system and identifying areas where

improvements are needed and gaps are apparent. Program and partner involvement for PA-DEW

is provided below:

Partner	Data Set	Purpose
Education	Basic education (K-12), career and technical education centers, community colleges and Perkins-funded institutions	Determine attributes, educational attainment, and skills of current and future labor market participants
Labor & Industry	New Hires	Track performance of new entrants and reentrants into the labor market to identify in demand education and skills
Labor & Industry	Wage records of employees	Track ongoing performance of workers in the labor market. Link back to education for analysis
Labor & Industry	Employment and wage data from employers	Track ongoing performance of industries, businesses and workers in the labor market. Partner with economic development for industries and companies of interest.
Labor & Industry	WRIS - Wage records of employees	Track ongoing performance of workers in the labor market in PA and other states
Labor & Industry	UC, Wagner-Peyser, Trade Adjustment Assistance	Determine cause of unemployment and implement corrective action based on occupation, education and skills
Labor & Industry Bureau of Workforce Development Partnership	Training & services for UC claimants –WIA Title I – Adult, Youth, Dislocated Worker	Determine effectiveness of further training or retraining
Labor & Industry Office of Vocational Rehabilitation	Training & services for OVR clients and job seekers	Determine effectiveness of further training or retraining
Public Welfare	Training & services for job seekers and welfare recipients, SNAP, TANF	Determine effectiveness of further training or retraining

3.2 Partnerships with State Education Agencies

L&I and PDE have collaborated for the last five years regarding the development and utilization of the state's High-Priority Occupations (HPO) model. In addition, L&I provides local occupational employment projections to PDE for use as one factor in the funding model of the Adult Basic and Literacy Education (ABLE) program. L&I also serves in PDE's CTE equipment grants review process. Finally, community colleges in Pennsylvania regularly work with L&I to focus their Economic Development Stipend program for trainings in regional HPOs. Two factors that drive the HPO process are the demand for jobs and the supply of qualified workers. Through the creation of PA-DEW, the strong collaboration between PDE and L&I will be further solidified with a joint mission of creating a real-time analysis of the potential supply of workers completing educational programs and ongoing measurement of the outcomes of such programs. PA-DEW will make possible Pennsylvania's first comprehensive supply-demand analysis that will be based on a model developed by the state of Florida. In preparation for the WDQI, L&I has reached out to the Florida workforce agency to understand their supply/demand model, which is recognized as the most comprehensive in the nation. With detailed information regarding training completers and their corresponding education, Pennsylvania will be able to gauge the success of programs in preparing individuals for the workplace and provide a realistic picture of the supply of workers for occupations in the state.

3.3 Partnerships with Research Universities or Other Research Entities

PDE houses PIMS, a part of our proposed federated data system, at the Enterprise Data Center located on the grounds of the Harrisburg State Hospital.

Regarding data system operations, the state has the internal capacity to develop and operate a longitudinal data system with initial vendor-assistance to develop the interface. Once developed,

it will be maintained and operated with current internal resources, making this project

economical to maintain and sustainable in the future.

L&I has a presence of analytical staff capable of conducting statistical and economic analysis.

While initially the state's focus is on utilizing current staff capacity to generate and automate

reports and perform research analyzing current programs, Pennsylvania's rich educational and

institutional heritage will be leveraged to conduct new and innovative research. L&I and PDE

plan to publicize the availability of this data on our websites and will reach out to the LWIBs,

institutions of higher education, and regional professional associations like the Pennsylvania

Economic Association, to ensure responsible but extensive use of PA-DEW. All data will

continue to reside in their respective agencies. Student level data will only be released pursuant

to PDE's *Data Access and Use* policy. Aggregate data will only be released pursuant to FERPA

regulations.

3.4 Partnerships with Additional State Agencies

Existing agency partnership with DPW will be leveraged in this project as well. Through a

14-year arrangement, L&I collects employer information under the New Hire law for all recently

hired and re-hired employees. While the primary objective of the program is to identify the

employment status of noncustodial parents not meeting child support obligations, this program

also provides real-time data on the labor market's hiring trends. New hire data will be the

primary dataset used to match individuals to education data and UC administrative data. L&I

also assists DPW in producing employment and earnings outcomes for participants in

Pennsylvania's Medical Assistance program. Furthermore, the New Directions Program and the

Employment Retention and Advancement Network (EARN) are integrated in CWDS,

Pennsylvania's one-stop system, and will be seamlessly integrated into the proposed longitudinal

database. The New Directions Program provides intensive placement services to welfare benefit recipients who are deemed work ready, while the EARN program addresses the needs of welfare recipients with barriers to employment.

Section 4: Database Design, Quality Assurance and Proposed Uses

Individual-level workforce, education, and New Hires data are stored in various databases throughout the commonwealth. Pennsylvania proposes to maintain the data in the existing databases and create linkages between the individual records to form the PA-DEW.

4.1 Personal Identifier

Pennsylvania's proposed federated approach is easier to maintain, less costly, and reduces privacy and security risks when compared to the creation of a new repository where data must be extracted from the various agency databases, moved, and separately stored. A key benefit to this approach is that it allows each partner agency to retain control over its respective data. Each agency can thus ensure that their data is accessed and used within the federal and state regulations to which they are bound. This federated data system will provide authorized users with an integrated view of the underlying databases through the use of a Business Intelligence (BI) /Decision Support tool. Standard reports and multidimensional cubes within the BI tool will query the multiple databases to answer research questions that cross agencies, programs and time. The data is compiled and presented to the user in a "single source" view. An ad-hoc reporting environment will be available and queries formed through this environment will likewise return data from the multiple databases in an integrated way.

Pennsylvania understands that disclosure of the data by any of the sharing partners has risk. The architecture described illustrates how the data will be kept intact. Further, data at-rest and data in-transit will be kept secure through advanced encryption technologies, depending upon the

data platform on which it is stored or presented. Although the optimal method for linking data is still being determined in the work being conducted under the USDE NCES IES SLDS grant, the proposed federated system is a solution that will provide the end-user analysts with access to individual-level de-identified data while protecting individual privacy and meeting the individual agency's state and federal regulations for data sharing.

Key to Pennsylvania's approach is the linking of data at the individual level. The state proposes to create a central directory of linked individuals that will be accessible to the PA-DEW solution. Since PDE does not collect social security numbers (SSNs), matching of individual records will be based on an algorithm that uses first name, last name, date of birth, and other demographic data elements that are available in both agencies' databases. Using the federated data system's Central Data Linking Directory, individual PII will not be revealed.

4.2 Data Quality Measures

The quality of the data in PIMS is of paramount importance to PDE. All staff in the Division of Data Quality have completed the USDE NCES Data Ethics Training Curriculum, a curriculum based on national best practices for high quality education data. Additionally, PDE utilizes a multi-step approach that includes both technology and process, as follows. The first step occurs early in the process as data is submitted into PIMS. Basic checks for data integrity are performed and feedback is provided to the submitter regarding errors that are encountered. Once the data has been submitted andaccepted into PIMS, validation and verification reports are generated and the submitter is alerted to additional data quality errors that are in need of correction before the data submission is considered final. When needed, PDE staff will work one-on-one with submitters that have specific data issues that arise and need individual attention. Once the data has passed the first two checks and corrections are submitted, the Chief School

Administrator of the submitting educational agency signs and submits an Accuracy Certification

Statement verifying the completeness and accuracy of the resulting data. PDE is continually

looking for ways to improve the quality of data and the timeliness of use. As such, PDE is

pursuing additional data quality initiatives, including a detailed training curriculum as outlined in

our latest USDE IES SLDS grant award.

The data collected by L&I is for the purpose of evaluating the various programs. Each

program has its own system in order to maintain quality and to validate the data. L&I utilizes

validation methods similar to those used by PDE in validating the PIMS data.

In a federated system as is proposed, the primary method of validation occurs within each

system. Each agency, through its Master Data Management strategy, individually strives to

reduce duplication and create consistency in service to our citizens. When data are viewed across

programs, the potential is expected to inspire efforts to streamline data from a state perspective,

further unifying separate agencies moving toward enhanced collaboration which will in turn

invoke efforts to improve data quality. However, through the WDQI grant the commonwealth

believes that we will be better able to cross-reference the data, both within the L&I systems and

between the L&I systems and PIMS. Further, by establishing a governance committee with

members from each department who will review the data and validation procedures, L&I and

PDE will be more equipped to develop and maintain improved validation techniques. The

committee will identify problematic data elements and suggest methods of remediation.

4.3 Longitudinal Data Scope

The federated data model has proven successful in Pennsylvania through the PIMS SLDS

project with early learning data stored in the DPW's Data Warehouse. Through this flexible,

scalable and easy-to-use solution, PDE data analysts use standard reports, an ad hoc reporting

environment and multi-dimensional cubes to analyze information on programs, providers, schools, staff, classrooms and children. This information is helping to improve early learning program services and the related student outcomes. Bringing together data from multiple state agencies will foster new ideas on data validation. Each agency, through its *Master Data Management* strategy, individually strives to reduce duplication and create consistency in service to our citizens.

The success of the federated data model will also be applied to link education data with workforce data. We anticipate utilizing this solution to incorporate data from WIA Title I, Wagner-Peyser Act, Trade Adjustment Assistance (TAA) (all in CWDS), and UI programs. The linkages will enable us to identify educational programs that may need to be supplemented by workforce training programs or workforce training that utilizes providers within the educational system, training gaps in both the educational and workforce system, and the supply of potential workers within both systems.

The details of the scope of the data are enumerated in the table contained in Section 3.1. All data housed by L&I as well as federal data (e.g. FEDES), can be linked using SSNs, however, PDE does not collect SSNs. PDE instead uses PAsecureID as a unique, anonymous student identifier. Therefore, since SSN matches are currently not possible, UC and new hire data will be used to match individuals with PDE files using basic demographic data available in all systems as outlined in 4.1. These datasets will also be used in analyzing outcomes associated with this project.

UC wage record information, FEDES, WRIS and new hire data, will form the foundation of employment and wage outcome data. The outcomes (detailed in other sections of the proposal) will include measures relating to placement rates, retention rates, wage levels and changes,

industry of placement, time elapsed until placement, etc. A fuller analysis of the types of research envisioned is contained in Section 4.5.

4.4 Security Measures

In a federated system, the primary responsibility of maintaining the security and confidentiality of the data rests on the agency that collects and houses the data. As detailed in Section 1.1, L&I has decades of experience with multiple databases in accomplishing this. Similarly PDE has done the same with education data and is vigilant in adhering to FERPA and all state and federal laws. The security issues therefore revolve around the linkages that are created between the two systems. As described in more detail in Section 1.1, we envision a federated virtual longitudinal data system where various data elements can be combined together and then de-identified. In this way the new data elements cannot be linked back to PII.

In addition, the governance committee, composed of members of both departments, will meet regularly to ensure that the data are properly de-identified and that security is maintained. Finally, L&I will leverage its partnership with a vendor who has experience in data integration to best utilize sharing of information among state agencies. Specifically, all data shared will be stored and maintained on agency host systems. This will allow security autonomy and data breach risk mitigation while still integrating valuable data assets across the commonwealth. As stated in section 1.1, the Commonwealth of Pennsylvania has employed strict policy on the use of its hardware, software and data assets. This policy, coupled with non-disclosure agreements within specific and related lines of business, creates security that can be measured against industry standards and best practices.

4.5 Planned Reports and Deliverables

The WDQI grant will be used to enhance the use of Pennsylvania's current data and capabilities in the areas of workforce, training and education. Currently, L&I generates many reports on the state of the labor force in Pennsylvania and performs evaluations of the outcomes of the various existing workforce programs. Similarly, PDE performs evaluations of its programs, but due to the lack of integration between workforce and education data, relies on survey data for employment information. PA-DEW will improve the quality of reporting used in analyses as well as expand the realm of questions that can be answered in an effort to improve training programs, student learning outcomes and labor market results. The improved data system will be beneficial to policy makers in evaluating effectiveness of current programs and in developing new programs. Specific areas of improvement include:

- Reports that allow Pennsylvania to improve evaluations of workforce programs from purely outcome evaluations to incorporate elements of process evaluations. Currently the state's main criterion of evaluating programs is post-employment status. PA-DEW will allow investigation of duration from post-program until employment as well. It will also be possible to statistically separate the impact of ability (as measured by proxies including GPA) from the impact of the program itself.

- Evaluating the impact of Perkins grants and measuring the return on investment of federal funds for adult literacy, as required by the WIA of 1998. The current system of data matching for adult literacy students will be less cumbersome with PA-DEW.

- Determine whether post-graduation students are employed in their field of study, and if so, the duration of time after graduation until that occurs.

- Analyzing the change in wages, pre and post education, of postsecondary students who have prior work histories.

- Comparing the employment and wage outcomes of postsecondary graduates to those who leave prior to graduation and evaluating wage benefits of postsecondary education relative to a high school diploma.

- Evaluating the interrelationships between workforce and education programs and analyzing the supply of potential workers within both systems.

PA-DEW will improve the ability of state agencies to analyze the efficacy of various programs involving education, training, and workforce. It will enable the commonwealth to partner with university researchers to create new knowledge specific to Pennsylvania. It is important to note that PA-DEW will be used by the commonwealth in the background of existing analytical research conducted by L&I using multiple Bureau of Labor Statistics (BLS) programs such as: Local Area Unemployment Statistics (LAUS), Quarterly Census of Employment and Wages (QCEW), Business Employment Dynamics (BED), Mass-Layoff Statistics (MLS) and Longitudinal Employer-Household Dynamics (LEHD). The new information created will also be used in Pennsylvania's Career Guide (annual publication), regional supply and demand conditions, Analysis of the Effects of the Minimum Wage, Profile of the Unemployed and numerous other regular publications produced by L&I on the state's labor market.

Section 5: Staffing Capacity

Staffing of the WDQI project will be accomplished with funding received from the grant in conjunction with current agency resources because the commonwealth realizes the tremendous value of this project. The project team will include a project director, program manager(s), a data analyst manager, database managers and data analysts. A *PA-DEW Steering Committee* will be formed to guide the powerful analysis that the project will generate and ensure the correct questions are answered for Pennsylvania's policy makers. The PA-DEW Steering Committee

will include Secretary Julia Hearthway of L&I, Secretary Ron Tomalis of PDE, Eric Kratz (L&I) and Erica Koser (PDE) of the Governor's Policy Office, David Malone, Chair, Governor's Workforce Investment Board, and a representative from an IHE selected by the Secretary of Education. The steering committee will meet quarterly prior to the Governor's Statewide Workforce Investment Board meetings to ensure connectivity to workforce and education efforts. PA-DEW's Steering Committee demonstrates strong intent to make the project successful. The core project team responsible for creation and implementation of PA-DEW consists of:

Project Director: Sue Mukherjee, Director at the Center for Workforce Information & Analysis (Pennsylvania's designated Employment Statistics Agency), will act as the project director for the WDQI project. Ms. Mukherjee will be responsible for the overall implementation of the grant and will serve as primary contact with USDOL on grant-related matters. Ms. Mukherjee has handled several large multi-agency projects for the commonwealth including serving as the state lead for Pennsylvania's Science, Technology, Engineering & Math (STEM) initiative, Prior Learning Assessment, Workforce Performance Management System, and Sector Cluster initiatives. Ms. Mukherjee will leverage her Economic Research Team, Workforce Information and Product Services Division for this project especially the expertise of Ed Legge, Division Chief, who will serve as the primary *Liaison* between L&I and PDE.

Project Managers, L&I: L&I's s project managers will include Randall Murphy and Edward Bowlen. Mr. Murphy will be responsible for meeting the grant objectives for CWIA and managing the work contracted under this project. He has over 35 years of experience in L&I and as a Division Chief in CWIA has worked extensively to develop the New Hire Reporting Program's longitudinal database (DM). Mr. Murphy is also the Chief Information Officer for Data Security in CWIA – this position interacts with L&I's Office of Legal Counsel to ensure

data agreements are appropriate. Ed Bowlen is one of the most innovative information technology leaders within L&I.

Project Managers, PDE: PDE Project Manager, Shara Bunis, will be responsible for managing all PDE tasks under this grant. Ms. Bunis, a Senior Project Manager from Shara Bunis Consulting, LLC, manages several PDE SLDS projects and has established collaborative relationships within L&I (CWIA specifically). Ms. Bunis' leadership will be critical in creation and advancement of PA-DEW.

5.1 Database Manager

Pennsylvania proposes to develop PA-DEW, a federated data system. Within this structure each department maintains its own data systems while an interface is created to link the various systems. As such, L&I and PDE will each maintain their own database manager(s) to oversee the operation, data quality, and security of their own data systems. The database managers from the two agencies will operate with partnership of the Project Director and Project Managers to ensure data emerging from the interface between the systems has a high level of integrity.

PDE Database Manager: David Ream will be responsible for ensuring the integrity of the data bridge being built to support the analysis of the linked data. Mr. Ream has 18 years information technology experience, including 11 years as an applications developer and seven years as the PIMS state longitudinal data system project director.

L&I Database Manager: Joe Sheridan will be responsible for ensuring the integrity of the data bridge being built to support the analysis of the linked data. Mr. Sheridan has over 20 years of information technology experience, including managing L&I's CWDS.

5.2 Senior Data Analysts

This project will have a lead data analyst involved in the structure design to identify necessary data components are captured and ensure the goal of longitudinal analysis is met. Timothy McElhinny will serve as this lead Senior Data Analyst for CWIA. As the current Manager of CWIA's Economic Research, Mr. McElhinny has a high level of knowledge, skills and abilities required to ensure the successful completion of the objectives of the grant proposal. He has twelve years of economic and statistical analysis experience and is highly proficient in the use of statistical analysis programs (SAS, SPSS, and SQL) for the extraction, manipulation, and interpretation of large and complex data sets. Mr. McElhinny will develop and recommend statistical and automated procedures, techniques, methodologies and tools for sampling the data contained in PA-DEW; and the analysis, interpretation, and dissemination of findings. Mr. Lee Siegel of L&I, a trained economist, will also serve as a Senior Data Analyst. Mr. Siegel will provide guidance regarding the automated reporting and will also have responsibility for utilizing PA-DEW to design and conduct economic and statistical analysis relating to the interaction among the variables. Mr. Siegel will be working under the direction of Mr. McElhinny.

Deb Rodrigues will be responsible for the integrity of the PIMS data quality, as well as ensuring all generated reports meet federal and state educational confidentiality requirements. Ms. Rodrigues, the Senior Data Analyst from PDE, has more than 20 years of educational experience, including eight years in educational data, five as the PIMS longitudinal data system assistant project manager.

Mike Dotts will be responsible for the integrity of Postsecondary PIMS data. Mr. Dotts has worked with postsecondary education policy and data analysis for 12 years in a variety of settings. Mr. Dotts has been involved in the PIMS project since it was expanded to include postsecondary data in 2009. Mr. Dotts has previously been involved in a pilot project with the 14

Pennsylvania community colleges and CWIA in which student records from the community colleges were matched to CWIA's wage record database to determine the wage gains of students completing a degree program.

5.3 Staff Qualifications

Specific qualifications are described in the positions listed above and resumes for select staff are included in the attachments to the Technical Proposal.

5.4 Staff Contributions to Data Sharing Partnerships

As part of the partnering and data sharing requirements of this project, all partners and staff involved in this project will be required to adhere to two commonwealth Information Technology Bulletins (ITBs) regarding data security: Encryption Standards for Data at Rest (ITB-SEC020) and Encryption Standards for Data in Transit (ITB-SEC031). Commonwealth staff involved in this project will also be required to adhere to their agencies and the commonwealth's confidentiality requirements pertaining to the data sets for which they may have access. Examples of such agreements include the UC Agent Agreement, BLS Agent Agreement, UCMS Access Agreement, and the CWDS Agreement.

PDE collection of data about students is compliant with the requirements of the federal Family Educational Rights and Privacy Act of 1974 (FERPA). Additionally, PDE has a Student Data Access & Use Policy. The purpose of this policy is to prescribe how data in PDE student unit record systems will be collected, maintained and disseminated in compliance with applicable federal and state laws. This policy applies to all organizational units and staff within the PDE, to LEAs and staff, to IPEs and staff, to authorized agents of PDE, LEAs or IPEs and any contractors, subcontractors and their agents, and to other individuals, agencies, and

organizations seeking to access or to utilize education records stored in PDE student unit record systems.

5.5 Employer

L&I's Center for Workforce Information & Analysis is the direct employer for Sue Mukherjee, Randall Murphy, Ed Legge, Timothy McElhinny and Lee Siegel. L&I's Office of Information Technology is the direct employer for Edward Bowlen, and Joe Sheridan. PDE is the direct employer for David Ream, Deb Rodrigues and Michael Dotts. PDE has contracted with Shara Bunis, a Senior Project Manager from Shara Bunis Consulting, LLC, Senior Project Managers in the Harrisburg, Pennsylvania area. A portion of staff compensation will be covered by sources outside of the WDQI grant.

Section 6: Other Data Linkages

Other data sources available in the longitudinal data system, PA-DEW, are as follows:

- L&I is a signatory on the *Wage Record Interchange System 2* (WRIS2) agreement which will enable capturing successful employment placements outside of Pennsylvania.

- L&I participates in the *Federal Employment Data Exchange System (FEDES)* which will enable capturing successful placements into federal employment (including the Office of Personnel Management, Department of Defense, Defense Manpower Data Center, and United States Postal Service).

- L&I is renewing its participation in an *Interstate Data Sharing Agreement* with five Mid-Atlantic states enabling sharing of wage records if a state is not a signatory to WRIS2.

Through experience as the department's Workforce Investment Act federal performance reporting entity, L&I's CWIA has strong experience in processing wage record requests through WRIS. Pennsylvania is therefore positioned to utilize the above resources to capture and

aggregate successful workforce placements outside of Pennsylvania and add to analysis of education programs. Additional data sources that will be tapped for this initiative come from the Pennsylvania DPW. L&I currently has an MOU with DPW to produce employment and earnings outcomes for participants in Pennsylvania's Medical Assistance Program. Through the New Directions Program, intensive placement services are provided to welfare benefit recipients who are deemed work ready. The EARN Program, which addresses the needs of welfare recipients with barriers to employment, is integrated into Pennsylvania's one-stop system. Both of these programs are part of the CWDS system which will be integrated into the longitudinal database as an enhancement to PA-DEW.

Conclusion

L&I and PDE maintain robust data systems. While being used for accountability, agency research, federal reporting, and client services, these systems have not been linked systematically. This proposed grant will formally establish a federated workforce and education data system for the state of Pennsylvania and will allow the state to: construct higher quality measures for workforce and education data; understand the impact of state and federal programs and policies; and facilitate more comprehensive research into the workforce and education systems. The proposed system, PA-DEW, will provide a basis for assembling data for Pennsylvania's key constituencies (students, workers and employers) to help grow the economy and create jobs. Key workforce and education data will provide a foundation Pennsylvania needs to advance, maintain and sustain prosperity and economic growth. Finally, in preparation of the PA-DEW, PDE will invest $182,000 of the 2009 SLDS grant to conduct an investigation of the data linking methodology and infrastructure PA-DEW will ultimately create.